Combining Demographic Analysis (DA) and ICM Results—An Overview

William R. Bell, Statistical Research Division, Bureau of the Census
October 29, 1997

1 The Problem: Differences Between DA and PES Results

DA has been used to evaluate census results for many years, as have coverage measurement surveys such as the 1990 Post Enumeration Survey (PES). DA can also be used to evaluate the PES results. Figure 1 shows such comparisons in terms of percent differences of DA and PES totals by age for Blacks and Nonblacks from the 1970, 1980, and 1990 censuses. The most notable features of these graphs are the large differences between DA and PES results for adult black males in all three censuses. Note that there is no evidence of underestimation for females (there is some suggestion of overestimation for Black females). The results of Figure 1 translate to a large difference between DA and PES sex ratios for Blacks as shown in Figure 2: There is also a difference in sex ratios, though much smaller, for Nonblacks. While the DA-PES differences shown in Figures 1 and 2 could be due to a variety of errors in both the DA and PES results, a leading explanation for the differences is *correlation bias* in the PES results for males.

Correlation bias can arise from two sources. The first is *heterogeneity* in probabilities of persons being included in the census and P-sample. Post-stratification for estimation attempts to address such heterogeneity, by specifying strata within which inclusion probabilities are hoped to be relatively homogeneous. The second potential source of correlation bias is a so-called *behavioral response*, where the act of being included in the census makes it more or less likely for a person to be included in the P-sample. Generally, correlation bias, when it exists, would be expected to be positive, leading to understimation by the usual Dual System Estimates (DSEs). Correlation bias can also be said to reflect a violation of the *independence* assumption underlying the usual DSEs.

2 Proposed Solution: Combine DA and ICM Results

Because of the historical observed differences between DA and PES results, research was done leading up to the 1990 census on ways to combine the two. Kirk Wolter

began research in this area, eventually publishing a paper on the subject (Wolter 1990). He developed two models to constrain the PES estimates to reproduce the DA *sex ratios* (# males / # females). These models only yielded results at the national level by age-race-sex, leaving open the question of how to combine DA with subnational (poststratum) PES estimates. I generalized Wolter's approach to address this problem (Bell 1993), developing a family of estimators that assume independence holds for females (usual DSEs are used), but that leads to modifications of the poststratum estimators for males such that, when aggregated to the national level within age-race groups, the DA sex ratios are maintained. These models can be thought of as taking the discrepancy between national PES aggregates for males and control totals for males formed by multiplying aggregated PES female estimates by the DA sex ratios, and allocating this discrepancy across the male poststrata. This is done within age-race groups. I considered four particular alternative estimators in my paper, all equally consistent with the PES and DA data, but all leading to different allocation schemes. Results for these four alternative models in terms of estimates of state population totals are shown in Figure 3, which also shows results from the usual DSEs. (For all the estimators shown, PES results from 357 poststrata are used.)

Still other models to combine DA and ICM results are possible. Some were proposed in Das Gupta and Robinson (1990) and others in the DAWG report discussed below. A further model leading to the simple result of multiplying all male adjustment factors within an age-race group by a constant to force agreement with the DA sex ratios is currently being written up.

One of the models I proposed in my paper (the model Little and Elliott (1997) call the "fixed relative risk" model) was selected for potential use in adjustment of the 1990 census. Howard Hogan and I presented it to the Undercount Steering Committee, which decided to delegate the decision on whether to combine DA and PES results to Charles Jones, Associate Director for the Decennial Census. Jones ultimately decided not to combine; his memo listing the reasons for this decision is included as an appendix to the DAWG report. Of course the issue eventually became moot when Secretary Mosbacher decided not to adjust the 1990 census.

In 1996 Ruth Ann Killion formed a Working Group on the Use of Demographic Analysis in Census 2000, known also as the Demographic Analysis Working Group or DAWG, for short. The report of this group (hereafter the DAWG report), issued May 6, 1996, contains some additional background and discussion of combining DA and ICM results. The main focus of the report is on research that would be desirable for making the decision on combining—consult the DAWG report for specifics.

3 Issue: Different Combining Models, Equally Consistent with the Data, Yield Different Results

Figure 2 illustrates an important issue with combining DA and ICM results. Namely, all four of the alternative DSEs used to produce the poststratum estimates which were aggregated to produce the state total results shown are equally consistent with the DA and PES data. Thus, the available data have no power to discriminate whether one of these models is better than any other, nor is any additional data that will permit such discrimination anticipated to become available for census 2000. (Essentially, what would be required is a "third system" from administrative records or another source that could be matched to both the census and ICM.) Technically, the problem is that all the models are "saturated," meaning that they require estimation of as many quantities as there are pieces of data, treating the DA sex ratios as one additional piece of data for each age-race group. The usual DSEs also come from a saturated model, though one that does not use the DA data.

Because of this problem, if the Census Bureau uses estimates that combine DA and ICM results, we could be in the position of having to defend the choice of model used for combining, without any data available to support this choice. That is, anyone disagreeing with our results could advocate use of a different combining model that would be equally defensible in regard to fit to the data, though of course they could not contend that the data supported their alternative model any more than ours. On the other hand, use of the original ICM estimates not combined with DA results will be open to the criticism that they are inconsistent with the DA data—i.e., they contain correlation bias, particularly for adult black males—and they are thus inferior to combined results. It is easier to make this argument for estimates of population totals than for population shares; the latter is more difficult for the same reason it is more difficult to establish superiority of usual DSE or any other ICM estimators to raw census counts or post-NRFU estimates for shares. Little and Elliott (1997) view the correlation bias of the usual DSEs as the more serious problem, and note that statistical problems that require modelling assumptions not checkable from the data arise in other areas. They note, for example, that adjustments for missing data must assume some sort of missing data model (e.g., the data are missing at random), typically without evidence to support such assumptions.

4 Policy Question: Can Direct ICM Estimates of State Totals be Changed?

The previous work on combining DA and ICM results has developed methods that modify the usual DSE poststratum adjustment factors in various ways so that the resulting estimators, when aggregated over poststrata to the national level within

age-race groups, agree with certain information from DA. Generally, DA sex ratios are the information used, though most of the methods would apply equally well to use of DA totals. In any case, since all male poststratum estimates are (potentially) altered to force agreement with DA nationally, generally by allocating DA-ICM discrepancies back to the poststrata, there is a sense in which, if combining were to be done in Census 2000, results from one state would have an impact on the combined estimates in another state. This impact would be indirect, coming from the relative contributions of the various states to the national DA-ICM discrepancy, rather than direct, as in 1990 when poststratum estimation used data from multiple states.

Das Gupta and Robinson have proposed some alternative methods of combining that constrain results to agree with direct ICM estimates of state totals obtained from the usual DSEs. To achieve agreement with national DA sex ratios, these, and any such methods subject to constraints of the direct ICM state estimates, must both *increase* estimates for males and *decrease* estimates for females. Several criticisms can be levied against such procedures: (1) DA totals suggest correlation bias is a problem for males but not females, so why change estimates for females? (2) Constraining combined results to agree with direct state totals from the usual DSEs, which DA suggests are underestimates due to correlation bias, is throwing out the primary benefit of combining. (3) We have not yet studied the effects of imposing the constraints of direct state estimates on combining with DA, and thus do not know whether it may lead to unreasonable results.

Before moving ahead to incorporate combining with DA into plans for ICM estimation in Census 2000, a decision is needed on whether such combining can be allowed to modify the direct state estimates obtained from the usual DSEs?

5 Draft Preliminary Report of Rod Little and Michael Elliott

Rod Little of the University of Michigan, with graduate assistant Michael Elliott, was contracted to review materials related to combining DA and ICM results, including the paper of Bell (1993) and the DAWG report. Their preliminary report (Little and Elliott 1997), currently in draft form, comments on several issues related to combining, some of which are mentioned elsewhere in this document. In general, they favor combining with DA on the grounds that the DA data supply information about census coverage that should not be ignored, and that suggests correlation bias in the usual DSEs. As noted above, they consider this more important than the issue that alternative combining models equally consistent with the data lead to different results. Having reviewed the alternative combining models proposed in Bell (1997) and the DAWG report, they prefer the second model proposed in Bell (1993), which

they refer to as the "fixed relative risk" (FRR) model. This model allocates, within age-race groups, the national DA-ICM discrepancy back to poststrata in proportion to the number of persons missed in the census in these poststrata, as estimated by the usual DSEs. Little and Elliott arrive at this preference through consideration of the following six principles they suggest for guiding model selection. (Note that these principles will not resolve the issue discussed in Section 3.)

1. PLAUSIBILITY—Does the model imply a plausible description of Census behavior?

2. FIT—Does the model fit the available data (not combining fails to fit the DA data)?

3. PREDICTION—Does the model yield reasonable predictions of key unobserved quantities, such as postratum undercount rates?

4. ICR INCLUSION—Does the model include the standard independence assumption capture-recapture model (usual DSEs) as a special case?

5. STABILITY—Given 1.–4., prefer models with lower variance.

6. CONSERVATISM—Given 1.–5., prefer models resulting in more conservative adjustments.

They rate the FRR model high on these principles. Das Gupta proposed a somewhat different set of criteria in Appendix B of the DAWG report, and arrived at a different preferred model. In selecting the FRR model for consideration for use in 1990, we went through a thought process very similar to that represented by the above list of principles, though we did not explicitly itemize things this way. In particular, Bob Fay used VPLX to compute variances of the four estimators from Bell (1993), and the results showed the FRR estimator to have generally lowest variance among the four considered.

As a personal reaction, having developed and applied the modelling framework in Bell (1993), I am less inclined than Little and Elliott towards the plausibility principle, as well as ICR inclusion. I would note that the fit principle essentially distinguishes just between combining and not combining. I am cautious about 4. given its highly subjective nature. I am most in agreement with 5. and 6. as guiding principles.

In addition to these comments, Little and Elliott suggest a more general statistical modelling approach to integrate both the original ICM DSE and the combining. Details remain to be worked out. Although a more integrated framework has some attractions, I would not expect the end results to be materially different from the approaches already considered. These considerations do lead Little and Elliott to suggest a somewhat different approach to the problem of negative cells, discussed next.

6 Additional Considerations—Limitations of DA and Negative Cells for DSE

The above discussion has centered on use of only sex ratios (by age-race groups) from DA. Use of DA totals was considered and rejected in 1990 because errors in DA estimates, particularly those arising from errors in estimates of undocumented immigration, are thought to be more serious for totals than for sex ratios (due to probably highly correlated errors for males and females). The DAWG report recommended research on using both sex ratios and age distributions from DA in combining, but no one, to my knowledge, has had time to thoroughly investigate this suggestion. Finally, while DA estimates for states and even metro areas have been developed for evaluating ICM and census results, Gregg Robinson does not feel these provide sufficiently reliable information to use in combining with direct ICM state estimates. The relatively high level of error in estimates of interstate migration is the concern here.

A problem that arose in the 1990 PES data was the occurrence of negative cells in the 2×2 tables used for DSE. This occurrend in roughly 1/3 of the 90 PES poststratum 2×2 tables. Negative cells arose when the P-sample weighted estimate of matches to the census for a poststratum exceeded the census count adjusted for erroneous enumerations. In short, the 2×2 table contains more matches than the census has people. This intuitively impossible situation can arise from sampling errors in the estimate of matches and erroneous enumerations (the latter is not very important, often when negative cells arose the matches exceeded the raw census count), and presumably from other errors (e.g., geocoding errors). The exact causes of this phenomenon are unknown, and thus the problem should be expected to recur in 2000.

The usual DSEs are not directly affected by negative cells, since the DSE uses only the P-sample match rate, erroneous enumeration rate, and census count (post-NRFU estimate), in forming the estimate, and not the individual cells of the 2×2 table. However, some of the alternative estimators that combine with DA do make use of individual cells. The extent of this problem depends on the specific form of the estimator; some are more affected by this problem than others. (Relative to the other models proposed in Bell (1993), the FRR model is the least affected by this problem.) In Bell (1993) I suggested modifying the 2×2 tables to eliminate negative cells before doing combining, and this is how the combining was carried out for evaluations of the 1990 census and PES. Little and Elliott (1997) suggest an empirical Bayes smoothing approach to address this problem. The DAWG report recommended research to determine why this problem occurs which, if successful, could suggest corrective actions that would benefit not just combined estimators, but DSEs uncombined with DA as well.

References

[1] Bell, William R. (1993) "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation," *Journal of the American Statistical Association*, 88, 1106-1118.

[2] Little, Rod and Michael Elliott (1997) "Census Adjustments Based on Combining Information from a Coverage Measurement Survey with Sex Ratios from Demographic Analysis," draft preliminary report on Activity 14.

[3] U.S. Bureau of the Census (1996) "Report of the Working Group on the Use of Demographic Analysis in Census 2000."

Figure 1. PES and DA Percent Differences for Three Census Years
100(DA - PES)/DA

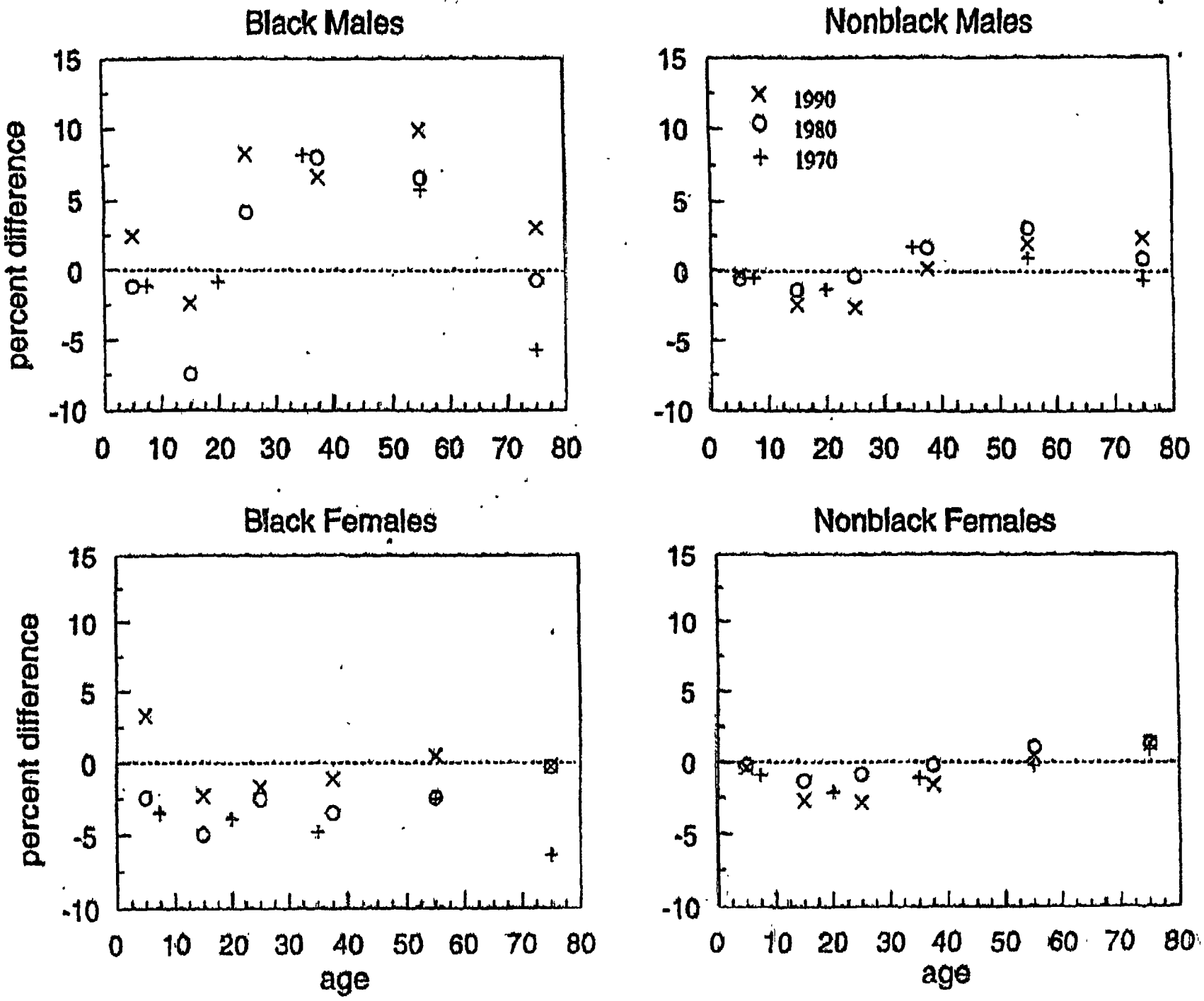
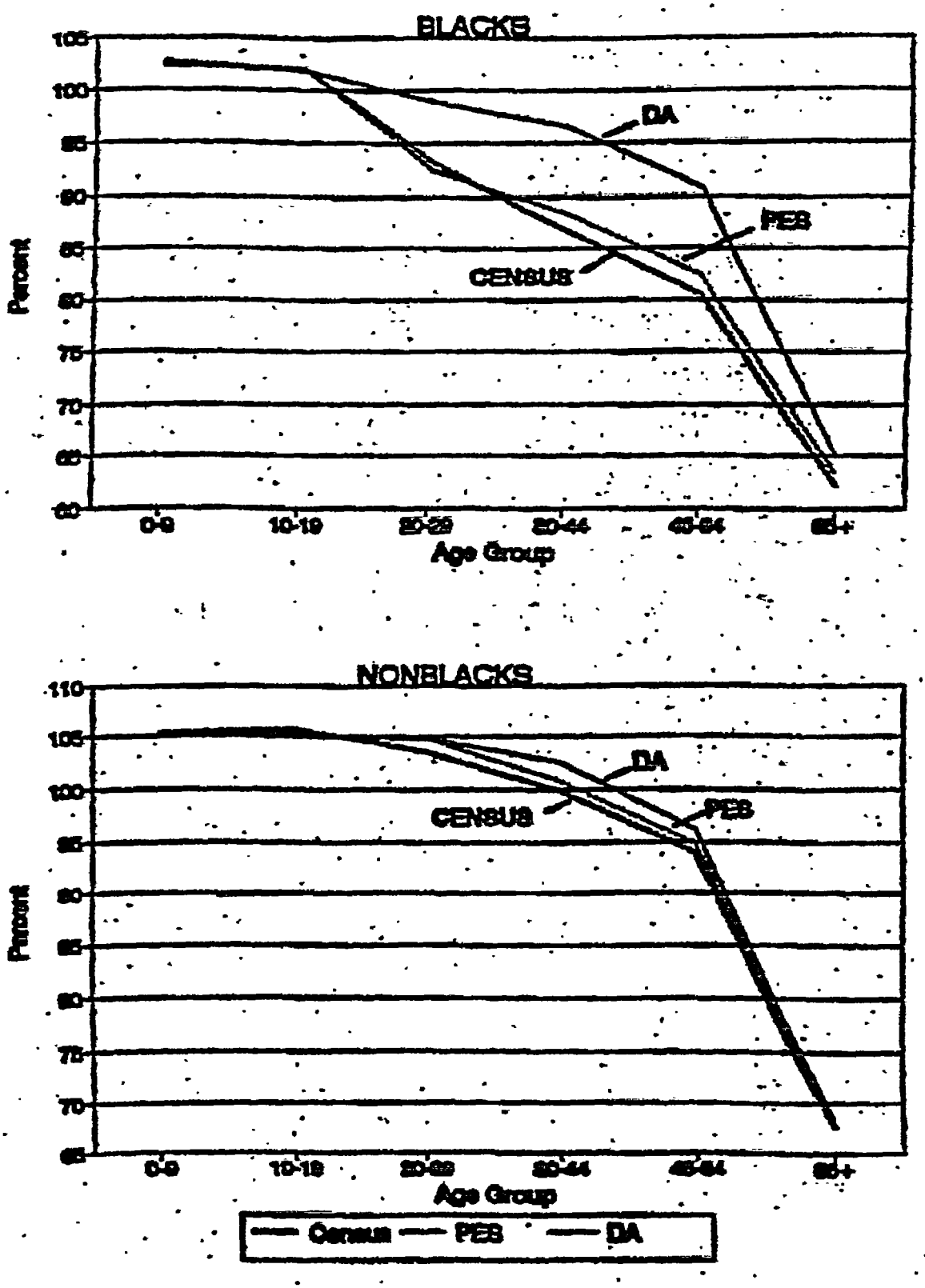


FIGURE 2

1990 Expected Sex Ratios: Comparison of Demographic Analysis & PES to 1990 Census



Note: Estimates pertain to the 1192 poststrata estimates. Sex ratios based on 357 poststrata estimates show the same patterns.

Figure 3 1990 State Undercount Rates from Alternative DSEs